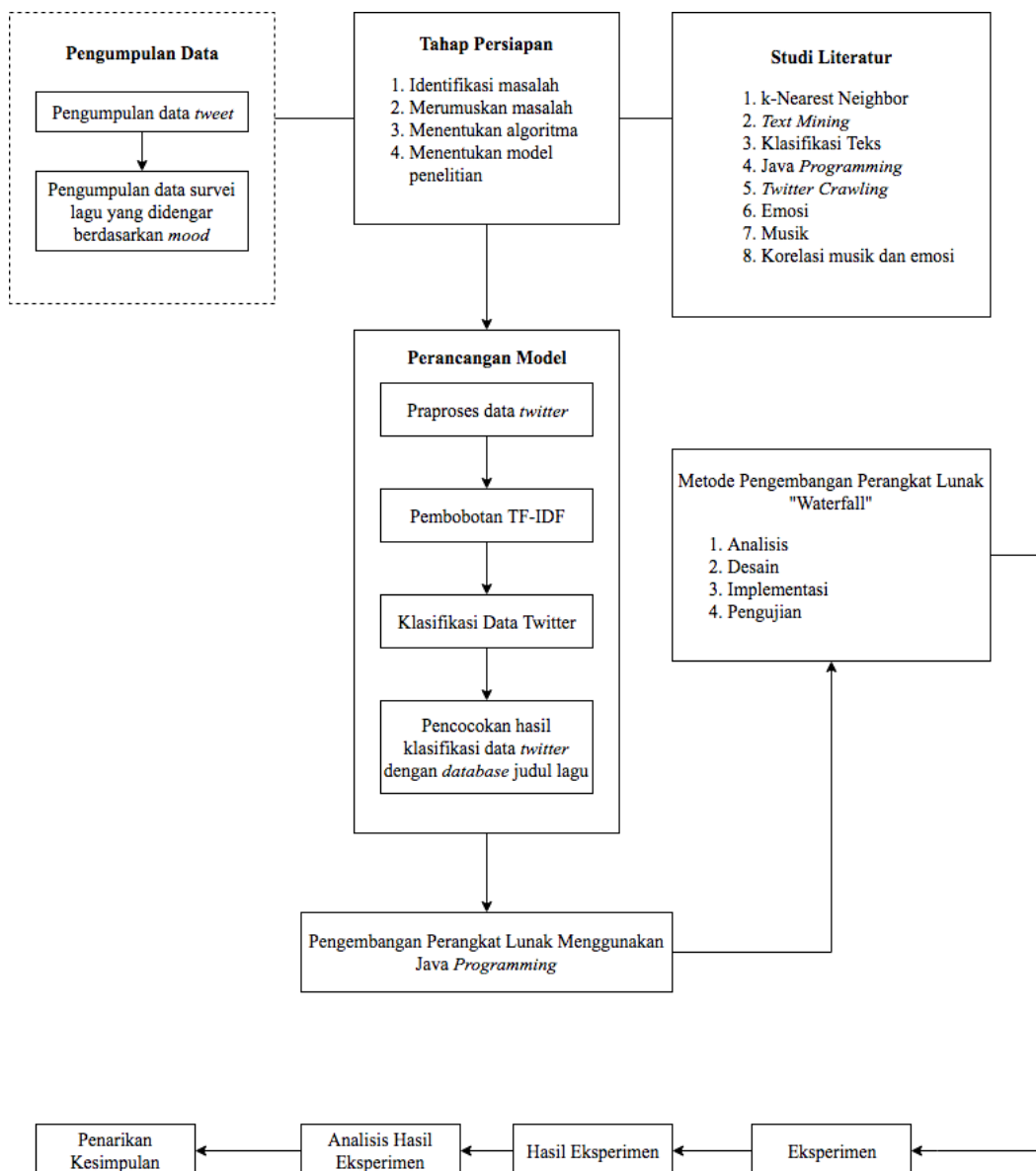


## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Desain Penelitian

Dalam penelitian ini desain penelitian merupakan salah satu tahapan yang dilakukan. Desain penelitian ini dijabarkan pada gambar berikut:



**Gambar 3.1** Desain Penelitian

**Nissa Aldawiyati, 2018**

IDENTIFIKASI EMOSI PADA TWEET DENGAN ALGORITMA K-NEAREST NEIGHBOR UNTUK REKOMENDASI JUDUL LAGU

Universitas Pendidikan Indonesia | [repository.upi.edu](https://repository.upi.edu) | [perpustakaan.upi.edu](https://perpustakaan.upi.edu)

Berikut penjabaran dari diagram sebelumnya:

### 1. Pengumpulan Data

Data dalam penelitian ini penulis mengumpulkan data tweet dari Twitter.com dan untuk mendapatkan data rekomendasi lagu, penulis membuat survei mengenai lagu yang didengar berdasarkan *mood*. Serta mengumpulkan data referensi judul lagu dari situs *allmusic.com*

### 2. Studi Literatur

Pada bagian ini merupakan proses mempelajari hal-hal yang terkait dengan penelitian, seperti K-Nearest Neighbor, konsep *text mining*, klasifikasi teks, Twitter *crawling*, emosi, musik, korelasi emosi dan musik serta Java *programming*.

### 3. Tahap Persiapan

Pada tahap ini terdapat empat bagian, yaitu:

- a. **Identifikasi masalah:** penelitian ini mengangkat permasalahan “*identifikasi emosi pada tweet dengan algoritma klasifikasi k-nearest neighbor untuk rekomendasi judul lagu*”.
- b. **Merumuskan masalah:** dari permasalahan yang telah dipaparkan maka dibuatlah rumusan masalah sebagai berikut: (i) bagaimana cara mengaplikasikan algoritma klasifikasi K-Nearest Neighbor untuk mengklasifikasikan emosi pada data *tweet* pengguna Twitter? (ii) bagaimana hasil akurasi, presisi, recall dan *F-Measure* pada setiap k dari klasifikasi *tweet* menggunakan algoritma k-Nearest Neighbor dapat merekomendasikan judul lagu sesuai dengan jenis emosi yang telah diklasifikasi? dan (iii) Pada nilai k seberapa algoritma K-Nearest Neighbor mendapatkan hasil akurasi, presisi, *recall* dan *F-Measure* terbaik?
- c. **Menentukan algoritma:** untuk memecahkan masalah yang telah diidentifikasi maka penulis menggunakan algoritma klasifikasi K-Nearest Neighbor (KNN).
- d. **Menentukan model penelitian:** pada bagian ini terdapat proses sebagai berikut:

### Praproses

Praroses yang akan dilakukan pada data Twitter sebagai berikut:

- a. **Cleansing:** proses pembersihan data teks dari hal-hal yang tidak diperlukan. Contoh pada kasus penelitian ini adalah menghapus simbol-simbol seperti *hashtag* (#), menghilangkan RT, *@username*, alamat *url* dan pembuangan *stopwords*.
- b. **Case Folding:** penghapusan angka dan tanda baca sehingga teks hanya berisi huruf latin dari a sampai z.
- c. **Tokenizing:** proses memecah data teks menjadi bagian per kata.

### Pembobotan TF-IDF

Pada tahap ini, data teks Twitter dikonversi menjadi bentuk vektor yang direpresenatiks dalam bentuk matriks dengan menggunakan *Vector Space Model* (VSM), kemudian setiap kata diberi bobot.

### Klasifikasi Data Twitter

*Tweet* akan diklasifikasikan ke dalam kelas yang sudah ditentukan sebelumnya. Pada proses ini digunakan algoritma KNN. Setelah data *tweet* dikelompokan maka data akan diuji kualitas dari model yang telah dihasilkan dengan menghitung nilai akurasi.

### Pencocokan Data

Hasil klasifikasi data Twitter yang telah dilakukan sebelumnya dicocokkan dengan dan data musik yang ada di *database*.

## 4. Pengembangan Perangkat Lunak

Penelitian ini mengembangkan perangkat lunak menggunakan model *waterfall* yang didalamnya terdapat beberapa tahap yang dilaksanakan seperti analisis kebutuhan, mendesain perangkat lunak yang akan dibangun, impelmentasi hasil desain, pengujian aplikasi yang telah dibangun dan juga hasil keluaran dari aplikasi tersebut.

## 5. Eksperimen

Pada bagian ini merupakan proses pengujian pada model yang telah dijelaskan sebelumnya. Model yang telah dirancang sebelumnya

diujikan dengan data baru yang diperoleh menggunakan algoritma KNN.

#### 6. Analisis Hasil Eksperimen

Pada bagian ini merupakan proses analisis dari hasil eksperimen yang telah di rancang pada bagian desain eksperimen. Analisis yang akan dilakukan adalah melakukan pencocokan antara data *tweet* dengan klasifikasi emosi dan lagu yang direkomendasikan.

#### 7. Penarikan Kesimpulan

Proses penarikan kesimpulan dari analisis hasil eksperimen yang dilakukan sebelumnya.

### 3.2 Metode Penelitian

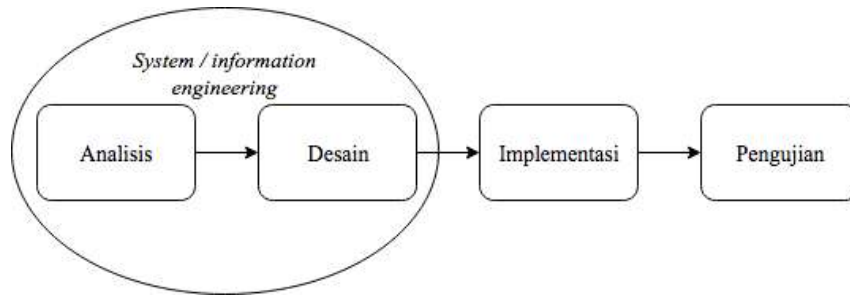
Pada penelitian ini, metode penelitian dibagi menjadi dua, yakni metode pengumpulan data dan metode pengembangan perangkat lunak. Berikut adalah penjelasan mengenai kedua metode yang digunakan:

#### 3.2.1 Metode Pengumpulan Data

Terdapat tiga data yang berbeda yang digunakan penulis dalam penelitian ini. Untuk mendapatkan data teks *tweet* pengguna diperoleh melalui *streaming* API Twitter. Sedangkan data kedua adalah data lagu yang diperdengarkan pada saat *mood* tertentu. Untuk mendapatkan data kedua, penulis membuat kuisisioner yang dibagikan secara *online* melalui layanan Google Form dan mengambil data judul lagu pada situs *allmusic.com*.

#### 3.2.2 Metode Pengembangan Perangkat Lunak

Pengembangan sistem dalam penelitian ini menggunakan model *linear sequential*. Model ini pula sering disebut dengan *classic life cycle* atau *waterfall*. Berikut adalah ilustrasi dari model *waterfall* (Pressman, 2001):



**Gambar 3.2** Model Pengembangan Perangkat Lunak "Waterfall

Tahap-tahap dalam model *waterfall* adalah sebagai berikut:

1. **Ruang lingkup rekayasa dan pemodelan sistem/informasi.** Perangkat lunak (*software*) merupakan bagian dari suatu sistem. Maka langkah pertama yang harus dilakukan adalah membangun semua syarat elemen dan kebutuhan sistem dan mengalokasikan bagian-bagian dari kebutuhan dan syarat ini menjadi sebuah perangkat lunak.
2. **Analisis.** Tahap pengumpulan data dan informasi. Pada tahap ini data *tweet* dan data musik dikumpulkan dan menentukan algoritma yang digunakan untuk melakukan identifikasi emosi.
3. **Desain.** Bertujuan untuk memberikan gambaran apa yang seharusnya dilakukan. Pada penelitian ini dibuat desain dari sistem yang akan dikembangkan untuk identifikasi emosi.
4. **Implementasi.** Tahap pembuatan perangkat lunak. Pada tahap ini implementasi program, *coding* dan integrasi modul-modul kecil.
5. **Pengujian.** tahap akhir dalam model *waterfall*.

### 3.3 Alat dan Bahan Penelitian

Alat yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Processor 2 GHz Intel® Core 2 Duo.
2. Memori 4GB 1333 MHz DDR3.
3. Monitor 13" NVIDIA GeForce 9400m 256MB.
4. Mouse dan keyboard

Adapun perangkat lunak yang digunakan:

1. Macintosh OS X Yosemite Version 10.10.5
2. Paket XAMPP 7.0.21-0

3. NetBeans IDE 8.2
4. Sublime Text 3
5. Library Twitter4J
6. Library Stemming Snowball

Sedangkan bahan penelitian diambil dari Twitter.